

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Novo Nordisk A/S
(B) STREET: Novo Alle
(C) CITY: Bagsvaerd
(E) COUNTRY: Denmark
(F) POSTAL CODE (ZIP): 2880
(G) TELEPHONE: +45 4444 8888
(H) TELEFAX: +45 4449 3256
(I) TELEX: 37304

(ii) TITLE OF INVENTION: Human Spasmolytic Polypeptide in Glycosylated Form

(iii) NUMBER OF SEQUENCES: 14

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Glu Lys Pro Ser Pro Cys Gln Cys Ser Arg Leu Ser Pro His Asn Arg
1 5 10 15

Thr Asn Cys Gly Phe Pro Gly Ile Thr Ser Asp Gln Cys Phe Asp Asn
20 25 30

29

Gly Cys Cys Phe Asp Ser Ser Val Thr Gly Val Pro Trp Cys Phe His
35 40 45

Pro Leu Pro Lys Gln Glu Ser Asp Gln Cys Val Met Glu Val Ser Asp
50 55 60

Arg Arg Asn Cys Gly Tyr Pro Gly Ile Ser Pro Glu Glu Cys Ala Ser
65 70 75 80

Arg Lys Cys Cys Phe Ser Asn Phe Ile Phe Glu Val Pro Trp Cys Phe
85 90 95

Phe Pro Asn Ser Val Glu Asp Cys His Tyr
100 105

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 563 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 77..553

(ix) **FEATURE:**

- (A) NAME/KEY: sig_peptide
(B) LOCATION: 77..235

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
(B) LOCATION: 236..553

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GAATTCATT CAAGAATAGT TCAAACAAGA AGATTACAAA CTATCAATT CATAACAAT 60

ATAAACGACC AAAAGA ATG AAG GCT GTT TTC TTG GTT TTG TCC TTG ATC 109
Met Lys Ala Val Phe Leu Val Leu Ser Leu Ile
-53 -50 -45

GGA TTC TGC TGG GCC CAA CCA GTC ACT GGC GAT GAA TCA TCT GTT GAG 157
Gly Phe Cys Trp Ala Gln Pro Val Thr Gly Asp Glu Ser Ser Val Glu
-40 -35 -30

30

ATT CCG GAA GAG TCT CTG ATC ATC GCT GAA AAC ACC ACT TTG GCT AAC Ile Pro Glu Glu Ser Leu Ile Ile Ala Glu Asn Thr Leu Ala Asn -25 -20 -15	205
GTC GCC ATG GCT GAG AGA TTG GAG AAG AGA GAG AAA CCC TCC CCC TGC Val Ala Met Ala Glu Arg Leu Glu Lys Arg Glu Lys Pro Ser Pro Cys -10 -5 1 5	253
CAG TGC TCC AGG CTG AGC CCC CAT AAC AGG ACG AAC TGC GGC TTC OCT Gln Cys Ser Arg Leu Ser Pro His Asn Arg Thr Asn Cys Gly Phe Pro 10 15 20	301
GGA ATC ACC AGT GAC CAG TGT TTT GAC AAT GGA TGC TGT TTC GAC TCC Gly Ile Thr Ser Asp Gln Cys Phe Asp Asn Gly Cys Cys Phe Asp Ser 25 30 35	349
AGT GTC ACT GGG GTC CCC TGG TGT TTC CAC CCC CTC CCA AAG CAA GAG Ser Val Thr Gly Val Pro Trp Cys Phe His Pro Leu Pro Lys Gln Glu 40 45 50	397
TCG GAT CAG TGC GTC ATG GAG GTC TCA GAC AGA AGA AAC TGT GGC TAC Ser Asp Gln Cys Val Met Glu Val Ser Asp Arg Arg Asn Cys Gly Tyr 55 60 65 70	445
CCG GGC ATC AGC CCC GAG GAA TGC GGC TCT CCG AAG TGC TGC TTC TCC Pro Gly Ile Ser Pro Glu Glu Cys Ala Ser Arg Lys Cys Cys Phe Ser 75 80 85	493
AAC TTC ATC TTT GAA GTG CCA TGG TGC TTC TTC CCG AAC TCT GTG GAA Asn Phe Ile Phe Glu Val Pro Trp Cys Phe Phe Pro Asn Ser Val Glu 90 95 100	541
GAC TGC CAT TAC TAAGTCTAGA Asp Cys His Tyr 105	563

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 159 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Lys Ala Val Phe Leu Val Leu Ser Leu Ile Gly Phe Cys Trp Ala
-53 -50 -45 -40

Gln Pro Val Thr Gly Asp Glu Ser Ser Val Glu Ile Pro Glu Glu Ser
-35 -30 -25

Leu Ile Ile Ala Glu Asn Thr Thr Leu Ala Asn Val Ala Met Ala Glu
-20 -15 -10

31

Arg Leu Glu Lys Arg Glu Lys Pro Ser Pro Cys Gln Cys Ser Arg Leu
 -5 1 5 10
 Ser Pro His Asn Arg Thr Asn Cys Gly Phe Pro Gly Ile Thr Ser Asp
 15 20 25
 Gln Cys Phe Asp Asn Gly Cys Cys Phe Asp Ser Ser Val Thr Gly Val
 30 35 40
 Pro Trp Cys Phe His Pro Leu Pro Lys Gln Glu Ser Asp Gln Cys Val
 45 50 55
 Met Glu Val Ser Asp Arg Arg Asn Cys Gly Tyr Pro Gly Ile Ser Pro
 60 65 70 75
 Glu Glu Cys Ala Ser Arg Lys Cys Cys Phe Ser Asn Phe Ile Phe Glu
 80 85 90
 Val Pro Trp Cys Phe Phe Pro Asn Ser Val Glu Asp Cys His Tyr
 95 100 105

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGCTGAGCCC CCATAACAG

19

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: synthetic

32

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGGAAACACC AGGGGAC

17

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GAGAAACCOCT CCCCCGCGA GTGCTCCAGG C

31

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCAGCCIGGA GCACTGGCAG GGGGAGGGTT TCTC

34

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
(A) ORGANISM: synthetic

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCTGAGAGAT TGGAGAAGAG AGAGAAACCC TCCCCCT

37

- (2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
(A) ORGANISM: synthetic

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TGOGTCATGG AGGCTCTC

17

- (2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
(A) ORGANISM: synthetic

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGCACCATGG CACTTCAAAG

20

- (2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

34

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GTCCCCCTGGT GTTTCACCC CCTCCCAAAG CAAGAGTOGG ATCAGTGOGT CATGGAGGTC

60

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 60 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TGAGACCTCC ATGAOCCACT GATCOGACTC TTGCTTTGGG AGGGGGTIGGA AACACCAGGG

60

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 46 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CATGGTGCTT CTTCCGAAC TCTGTGGAAG ACTGCCATTA CTAAGT

46

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 46 base pairs

35

- (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: synthetic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CTAGACTTAG TAATGGCAGT CTTCCACAGA GTTCGGGAAG AAGCAC

46

Sim
BI
cont